RESEARCH ARTICLE

Key Points:
- Develop ensemble CME forecast system and pertinent analytic and graphical tools
- Nonchaotic propagation implies different tactics than for terrestrial weather
- Establishes solid theoretical basis for forecast and analysis of real events

Correspondence to:
V. J. Pizzo,
Vic.Pizzo@noaa.gov

Theoretical basis for operational ensemble forecasting of coronal mass ejections

V. J. Pizzo[1], C. de Koning[1,2], M. Cash[1,2], G. Millward[1,2], D. A. Biesecker[1], L. Puga[1], M. Codrescu[1], and D. Odstrcil[2,3]

[1]NOAA/Space Weather Prediction Center, Boulder, Colorado, USA, [2]CU/CIRES, Boulder, Colorado, USA, [3]School of Physics, Astronomy, and Computational Data Sciences, George Mason University, Fairfax, Virginia, USA

Abstract We lay out the theoretical underpinnings for the application of the Wang-Sheeley-Arge-Enlil modeling system to ensemble forecasting of coronal mass ejections (CMEs) in an operational environment. In such models, there is no magnetic cloud component, so our results pertain only to CME front properties, such as transit time to Earth. Within this framework, we find no evidence that the propagation is chaotic, and therefore, CME forecasting calls for different tactics than employed for terrestrial weather or hurricane forecasting. We explore a broad range of CME cone inputs and ambient states to flesh out differing CME evolutionary behavior in the various dynamical domains (e.g., large, fast CMEs launched into a slow ambient, and the converse; plus numerous permutations in between). CME propagation in both uniform and highly structured ambient flows is considered to assess how much the solar wind background affects the CME front properties at 1 AU. Graphical and analytic tools pertinent to an ensemble approach are developed to enable uncertainties in forecasting CME impact at Earth to be realistically estimated. We discuss how uncertainties in CME pointing relative to the Sun-Earth line affects the reliability of a forecast and how glancing blows become an issue for CME off-points greater than about the half width of the estimated input CME. While the basic results appear consistent with established impressions of CME behavior, the next step is to use existing records of well-observed CMEs at both Sun and Earth to verify that real events appear to follow the systematic tendencies presented in this study.

1. Introduction

Ensemble modeling has played an important role in terrestrial weather forecasting and other applications for some time. A well-known example is in the area of hurricane prediction, where the familiar plots of likely storm paths are generated either from a variety of model runs with slightly different initial conditions (as in Figure 1, top), or from a series of runs with different models, each with its own set of detailed descriptions of physical processes, or from both. The bulk of the work in hurricane prediction has been developed with the realization that the underlying physical system—terrestrial weather—is dominated by such a plethora of sources and sinks of energy and momentum that it must be characterized as "chaotic." This means that starting from some initial state, the future state of the system can be forecast only for a short period, beyond which error due to uncertainties in all the myriad competing processes grows exponentially. That is, the output for simulations whose initial states differ only infinitesimally can diverge wildly after a certain time, or simulations with quite different initial conditions may all coalesce to some "attractor" state.

Not all terrestrial physical systems evolve chaotically. In particular, the Center for Tsunami Research at the National Centers for Environmental Prediction (NCEP) generates tsunami forecasts under entirely different assumptions—namely, that the propagation of tsunamis is contingent upon a small set of knowable factors, such as the magnitude and location of the initial earthquake pulse, the geometry of the ocean basins, and the wave properties of water. Specialized models can then be run to forecast the impact of the tsunami in any area of the globe, given the inputs mentioned above. Since each such computer run can take longer than for the tsunami to propagate to shore, a vast array of simulations is computed in advance, covering a broad range of inputs of varying magnitudes and locations likely to be encountered sometime in the future. Then, once an undersea event is detected (e.g., Figure 1, bottom, the great tsunami of 2006 along the coast of Japan), the propagation forecast is made by using a lookup table, with the prediction at any location being made on the basis of the run or runs most nearly matching the properties of the current event [*Gubler et al.*, 2013]. All of this hinges upon the fact that the propagation of tsunamis in the oceans is a "nonchaotic"
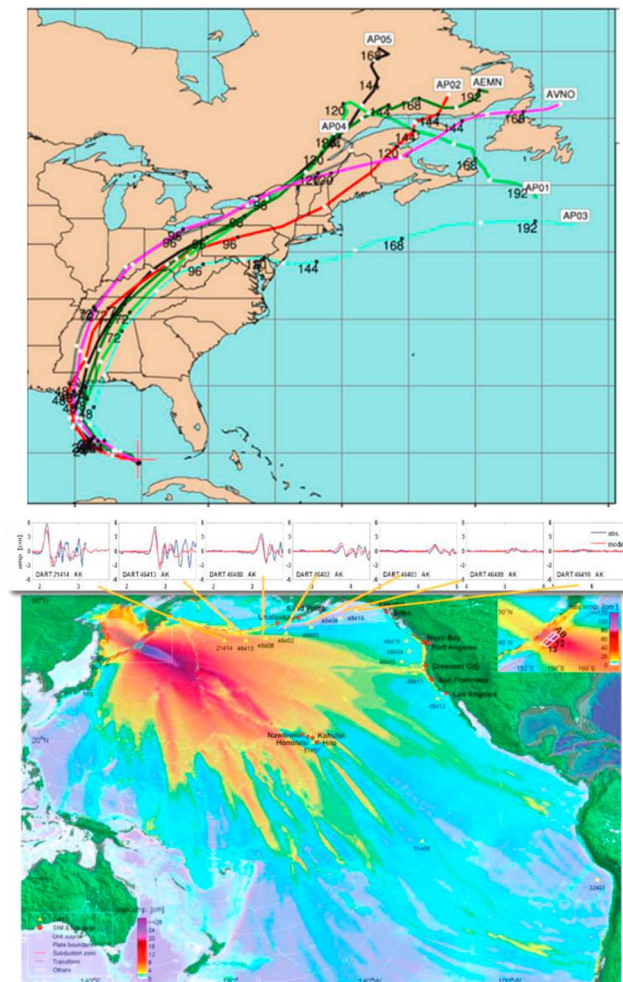
Figure 1. Examples of chaotic (e.g., hurricane) and nonchaotic (e.g., tsunami) model forecasting in operations at NCEP. (top) Advance Global Forecast System ensemble forecast tracks for Hurricane Katrina (courtesy https://www.e-education.psu.edu/worldofweather/s12.html); (bottom) a tsunami propagation forecast comparison with buoy data for the 15 November 2006 Kuril tsunami (http://nctr.pmel.noaa.gov/Jpg/kuril06-dart.jpg).

physical system, wherein the outputs are directly—and repeatedly—related to the inputs and small changes to inputs result in consistent, corresponding departures in outputs. This is in stark contrast to the chaotic nature of terrestrial weather, for which such an approach would be useless on all but the very shortest time scales.

It is useful to understand which kind of physical system is involved in forecasting the timing and impact at Earth of coronal mass ejections (CMEs) detected in the solar corona. Indeed, as described above, very different tactics are mandated depending upon whether the forecasting is undertaken under chaotic versus nonchaotic regimes.

Ensemble modeling of CME propagation has historically seen a bit of both approaches. Accurate prediction of ambient solar wind properties lying in the path of CMEs has long been recognized as an essential component of CME forecasting [e.g., *Case et al.*, 2008; *Shen et al.*, 2014]. (The term "ambient" in this paper means the pre-CME background solar wind, which may be either uniform or structured, and we consider both cases here.) More recently, variations in ambient flow forecasts stemming from model specifics, from differences in input parameterizations, and from varying numerical grid resolutions have been studied using ensembles [*Riley et al.*, 2013]. Ensembles also play an integral role in the Air Force Data Assimilative Photospheric Flux Transport approach to improving the Wang-Sheeley-Arge (WSA) solar wind modeling system [*Arge et al.*, 2003, 2013; *Linker et al.*, 2013; *Hickmann et al.*, 2015]. Ensemble methods have also been applied to the study of specific CME events [*Lee et al.*, 2013, 2015] using what amounts to a nonchaotic approach, while *Emmons et al.* [2013] employed random-member ensembles more in line with chaos methods. Recently, *Mays et al.* [2015] have reported on the application of ensemble concepts to CME forecasting in a near-real-time (NRT) environment. Finally, it is worth noting that ensemble modeling efforts are also being undertaken in the broader context of space weather studies [e.g., *Godinez and Koller*, 2012; *Murray et al.*, 2015].

What has been particularly lacking in the CME ensemble studies to date, though, is a comprehensive, step-by-step analysis of the underlying physics of the combined ambient flow (e.g., WSA) and CME propagation (e.g., Enlil) modeling approach used in the operational environment [*Pizzo et al.*, 2011]. There is a need to establish across a broad spectrum of assumed CME properties (speed, size, direction, etc.) and ambient states (i.e., global corotating interaction region (CIR) structure) just how these inputs relate systematically—if they do—to what is seen at Earth or any other observing point. We will also be able to see whether there is any evidence of chaotic behavior, as laid out above. The *Riley et al.* [2013] study suggests that there will not be, and, if so, that will guide the approach we take toward constructing a formalism for quantitative assessment of operational model forecasts.

Table 1. Physical Variables for the Enlil "a5b1" Uniform Steady State at the Inner Boundary (Left) and at 1 AU (Right), for a Range of Input Conditions

	$R = 21.5\ R_S$			$R = 1$ AU	
V (km/s)	n (cm^{-3})	T (K°K)	V (km/s)	n (cm^{-3})	T (K°K)
250.0	1,176.0	255.1	256.9	11.27	11,599.0
300.0	816.7	367.4	319.6	7.262	16,321.0
350.0	600.0	500.0	380.6	5.441	21,925.0
400.0	459.4	653.1	440.6	4.114	28,401.0
450.0	363.0	826.5	499.9	3.223	3,574.07
500.0	294.0	1,020.4	558.9	2.595	43,961.0
550.0	243.0	1,234.7	617.5	2.136	53,042.0
600.0	204.2	1,469.4	675.9	1.789	62,988.0
650.0	174.0	1,724.5	734.1	1.520	73,801.0

Toward that end, we enunciate a succinct logic for minimizing the number of ensemble members needed for operational applications, and we develop a set of well-conceived, readily understood analysis tools for quantifying and ultimately reducing error in CME forecasts in the longer term. Finally, armed with that broader understanding, we will be in a position to develop meaningful and feasible tests of the efficacy of the forecast system by data-mining collections of past events for which adequate observational information is available. On that basis, we can then evaluate the performance of the system going forward against NRT events, for which we do not know the outcome in advance.

These latter two aspects, data-mining and NRT applications, will be deferred to a later paper. However, we do introduce some ideas on how we may best proceed in those directions and how we may possibly simplify the ensemble modeling tactics that are ultimately to be applied to NRT CMEs.

2. Modeling System and Approach

For all these simulations, we use the coupled WSA-Enlil model (WSA V2.6 and Enlil V2.7e [*Odstrcil et al.*, 2005]). Enlil V2.7e is quite similar to that currently used in Space Weather Prediction Center (SWPC) operations (V2.6.2) and elsewhere. The main difference between the two is a small quantitative variation in the way that the ambient solar wind is specified. Since the flow density and temperature at the Enlil inner boundary (located at $R_0 = 21.5$ solar radii) cannot be measured directly, these quantities are specified so as to mimic the mean values seen at Earth orbit. The range of values that can meet this general requirement is relatively broad, and various combinations have been employed, some providing a good match to observed parameter ranges during one epoch, others during another. Within the WSA-Enlil system, these are distinguished by a simple naming convention that is invoked at compilation. For example, the "a3b2" specification is the default in the SWPC operational version, whereas "a5b1" is the default in V2.7.e. Table 1 provides a listing of the thermodynamic specifications and their corresponding 1 AU values for the "a5b1" variant used in this study. (In general, "a5b1" provides a better parameterization for active solar conditions, but the overall effect upon corotating stream and CME propagation is minor, as compared to all the other uncertainties and approximations involved in such simulations.)

Commensurate with the considerable uncertainties in CME inputs and the unavoidable crudeness of characterizations of CME structure, it suffices to conduct ensemble studies with the WSA-Enlil system run in the 4° × 4° "coarse" angular resolution mode. (Moreover, practical computational resource considerations dictate that ensemble studies be run as efficiently as possible, which also implies coarse grid. SWPC operational simulations, which are run in "medium" 2° × 2° resolution over about the same radial domain, take over an hour, whereas coarse-grid solutions can be obtained in a minute or two.) In the radial direction 240 grid nodes cover the range from the 0.1 AU inner boundary out to 1.7 AU. Simulations are run with 24 processors on the NCEP development machine "Zeus" (see https://nesccdocs.rdhpcs.noaa.gov). Although these coarse-grid solutions are somewhat smooth (not featuring substructure likely to be associated with CMEs propagating in a structured medium), gross properties such as arrival time and relative CME strength (as measured by velocity and density jumps across the front) are nonetheless representative of real CMEs. Moreover, we find that in practice these basic properties differ insignificantly from those obtained by the SWPC operational model at medium grid resolution.
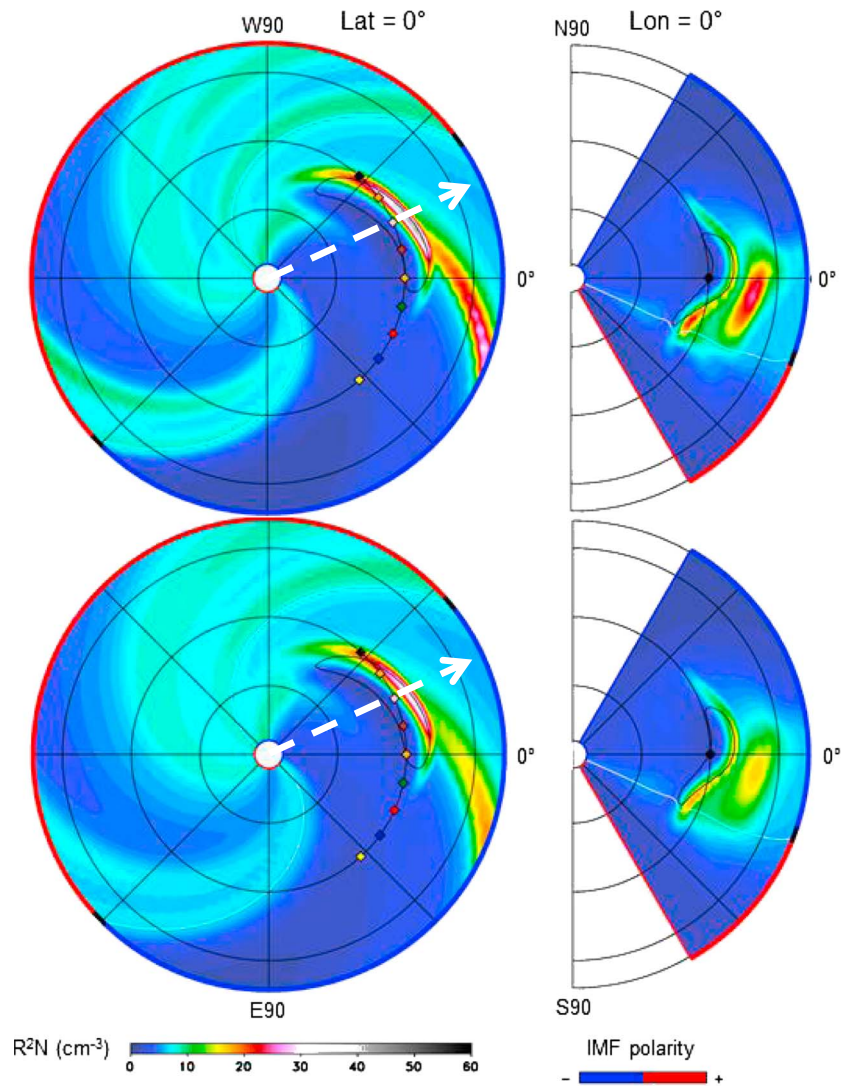
Figure 2. Comparison of mass density distributions in two CME simulations, (top row) one using a purely HD ambient flow, (bottom row) the other including the spiral IMF in the ambient. (left column) The solution in the solar equatorial plane and (right column) meridional slice along $\varphi = 0$. The locus of the injected HD CME mass is indicated by the fine black line. Both simulations were run in coarse-grid mode, with 240 evenly spaced radial steps between the inner boundary at 21.5 Rs and the outer boundary at 1.7 AU and an angular grid spacing of $\Delta\lambda = \Delta\varphi = 4°$.

For this study we restrict attention to the default Enlil CME cone inputs, which include the estimated speed V_{CME}, the north-south (λ) and east-west (φ) launch directions (latitude and longitude of the CME centerline, in Heliocentric Earth Equatorial coordinates), and the angular half width $\omega_{1/2}$ (here, assuming a circular cross section). For our hypothetical CMEs, the final operational input—the time of arrival of the nose of the CME at the Enlil inner boundary R_0—is arbitrary. The injected CME material is modeled as a sphere of uniform density crossing R_0 at the rate V_{CME}, with the density being 4 times that of the a5b1 reference density (see Table 1). This results in a direct relationship between total mass injected and angular width, functionally similar to, but quantitatively different from, that reported by *Gopalswamy et al.* [2005], for instance.

A number of additional modifications to the cone inputs are provided for in the Enlil code, but we defer consideration of those to a later paper. Finally, we note that magnetic cloud content is not in this model, and it can be expected that its inclusion would affect the internal CME evolution and its impact upon Earth (duration and strength of a geo-event) quite substantially. However, that is outside the scope of this study, which is restricted to properties near the leading edge of the CME, where the evolution is dominated by hydrodynamic considerations.
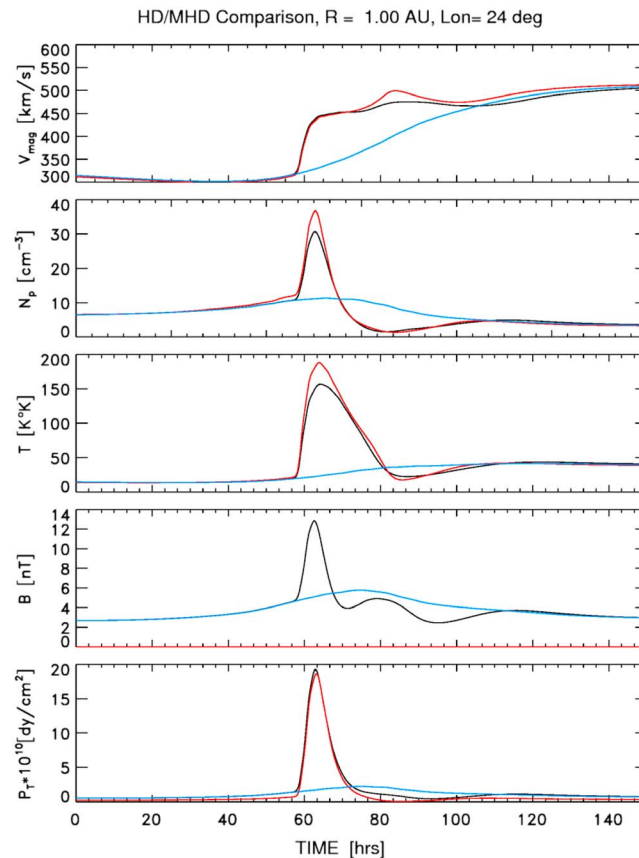
Figure 3. Overlay of time profiles of several flow parameters along a line near the center of the CME in Figure 2. The MHD solution with spiral IMF is shown in black, the pure HD solution is red, and, for reference, the undisturbed MHD ambient is in blue. The gas pressure in the HD solution is higher than in the MHD case because the IMF pressure in the latter helps oppose the pileup of momentum at the leading edge of the CME. The amplitudes of all relevant variables in the two CME solutions are very similar, and the arrival time is affected hardly at all.

Our approach is to study the impacts of variations in CME cone inputs—speed, width, and direction—upon CME transit time to 1 AU in a comprehensive way, including explicit consideration of the effects of varying ambient speeds upon the propagation time in conjunction with the different cone inputs. This is a purely hypothetical study intended to scope out the range of behaviors and responses over a representative sample of cone inputs and ambient conditions. This will lay the groundwork for quantitative follow-on studies involving real CMEs that have previously been modeled and forecast in the SWPC operations center.

Here we consider two broad classes of single-CME evolution: (1) CMEs launched into uniform hydrodynamic (HD) backgrounds, and (2) CMEs launched into the midst of a substantial tilted-dipole stream structure. The uniform HD background study provides a useful introduction into the basics of ensemble CME composition, gross dynamic evolution, and the specifics of analysis techniques. This sets the context for handling and understanding CMEs cast into a more realistic tilted-dipole global stream configuration in the second stage of this study.

It may come as a surprise to some, but for our purposes there is no loss of generality in neglecting the ambient spiral interplanetary magnetic field (IMF). That is, a 3-D MHD Enlil simulation of a CME interacting with ambient stream structure in the inner heliosphere, as driven by typical WSA map information and cone inputs, will exhibit only minimal differences at 1 AU as compared to a 3-D HD Enlil simulation (i.e., where the magnetic field $|B|$ has been set to zero everywhere) driven by the same velocity, density, and temperature variations at 21.5 R_S. In terms of the effect upon CMEs—in particular, fast CMEs—interacting with the ambient stream structure, there is essentially *negligible* difference in CME front properties such as arrival time and overall velocity jumps across the CME fronts, as compared with all the very real uncertainties in the inputs. The underlying reason for this is that the distribution of input radial momentum flux dominates the evolution in the interplanetary medium, with the thermal and magnetic pressures playing a secondary role [*Pizzo*, 1980].

To illustrate, Figure 2 (top row) shows an Enlil time slice pair of plots for an HD simulation wherein a moderate-speed CME has been launched into the back end of an ambient corotating stream front driven by an actual WSA map (the one included in the V2.7e release of the Enlil software). The input cone parameters were $V_{CME} = 800$ km/s, $\lambda_{CME} = 0°$, $\varphi_{CME} = 30°$, and the full angular width $\Omega_{CME} = 60°$. The plots present an overview of the resulting structure in the inner heliosphere, about 3 days after CME launch. Figure 2 (top, left) is a color-coded representation of the density in the solar equatorial plane, and Figure 2 (top, right) is a north-south cut along a plane passing through Earth (the gold diamond at 1 AU and 0° longitude). Figure 2 (bottom row) shows the equivalent MHD simulation, otherwise the same as the one at the top but including a typical spiral interplanetary magnetic field (IMF). Comparing the two sets of plots, it can be
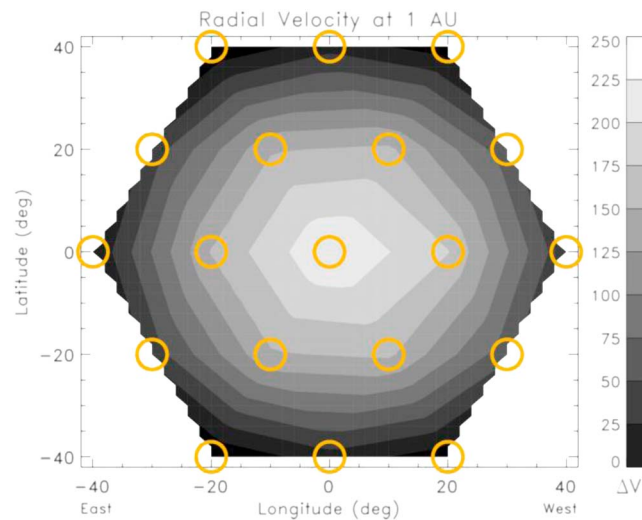
Figure 4. Latitude-longitude contour map of 1 AU radial velocity amplitude across a sample moderate-strength CME launched into a uniform HD solar wind background. The contours indicate the velocity jump relative to the pre-CME background, and the yellow circles denote the loci of radial lines along which the properties of the CME are measured for purposes of analysis.

seen that the density peaks in the HD CME front are a bit higher and sharper than in the MHD case, but the locations and timing of both ambient and CME features are almost indistinguishable.

To appreciate more keenly how little the inclusion of a spiral IMF affects the CME propagation, consider Figure 3, which shows 1 AU time series plots along $\varphi = 24°$, in the direction of the dashed white arrow in the equatorial plane in Figure 2 (left column). The first panel of Figure 3 presents overplots of HD (red) and MHD (black) solar wind velocity, the second panel number density, the third panel temperature, the fourth panel magnetic field intensity, and the fifth panel total pressure. The blue curves represent the undisturbed MHD ambient solution (no CME).

Experiments with MHD versus HD solutions across a range of CME inputs like that explored in section 4.1, below, confirm the generality of the conclusion that use of a purely HD versus an MHD model calculation causes negligible change in the bulk dynamics. Indeed, it is the case that the higher the CME speed—and the greater the potential for geospace disruption—the more rigorous is this finding.

3. Quantifying and Organizing Results—A Sample Ensemble

We illustrate the process of composing, running, and analyzing CME ensemble simulations in the simplest of all scenarios: an ensemble of CMEs is launched into a uniform HD ambient flow, set to a moderate 350 km/s at R_0. We build our ensemble about a base CME of moderate speed and angular width {$V_{CME} = 800$ km/s, $\omega_{1/2} = 35°$} as follows.

First, we establish the spread in cone input variables about a base member. Given all the inherent uncertainties in cone specification derived from coronagraph observations, it is reasonable to start with an ensemble composed of values bracketing some base member. In terms of V_{CME}, we select a spread of ±200 km/s, for $\omega_{1/2}$ a spread of ±10°, and for λ_{CME} and φ_{CME} ±10°. These values are representative only and are based on experience in fitting cones to real CMEs; they are, moreover, gauged to provide enough variation in 1 AU CME properties to support systematic investigation of ensemble properties, as developed below. For just this simple ensemble, 81 members (=3^4) would thus be needed to cover all potential combinations. Since each coarse-grid solution requires about 30 s of CPU time on a modern parallel supercomputer like Zeus, a full ensemble would entail about 40 min of computing in the optimum case. In itself, that poses no practical issue, but when one contemplates running hundreds of such ensembles (as undertaken in this study), the computational load does become a consideration. Regardless, a far more compelling motivation for minimizing the number of members in each ensemble is the task of extracting useful information from the large amount of output. Although it is possible to automate that process to some extent (see below), there is a huge advantage in not complicating the problem anymore than absolutely necessary. Also, it must always be kept in mind that the objective is to draw relatively general conclusions from a system that is built upon crude inputs that are coarsely modeled.

For each ensemble member, certain information must be efficiently extracted and saved in an organized way for subsequent analysis. Since we are dealing with structural variations across a broad front, we gather data on such features as local transit time (TT) and velocity and density amplitudes (ΔV and Δn) at a variety of observing points, as illustrated in Figure 4. Here we view a representation of the velocity distribution in our
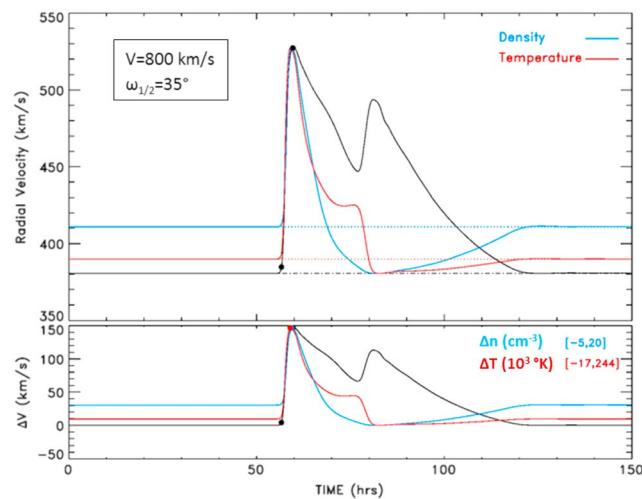
Figure 5. Illustration of automated methodology for collecting parameter variations across CME fronts used in the study. (top) Raw parameter values as a function of time from CME launch, with the scale at left referring to the velocity (black), while the density (blue) and temperature (red) are arbitrarily scaled to the velocity amplitude. Dotted lines indicate the undisturbed ambient values. (bottom) The parameter amplitudes relative to the pre-CME background, with the Δn and ΔT ranges as indicated. The amplitude data and timings (colored dots) are collected and stored in comma-separated values (CSV) files for subsequent processing and display. For CMEs propagating in structured backgrounds, the location of the several parameter peaks may vary slightly.

base CME, with the contours depicting the distribution of maximum velocity jump across the CME front. It is to be emphasized that the contours are not a snapshot at a given time but rather are created (by irregular grid triangulation) by sampling the velocity jump (relative to the pre-CME background) at 1 AU at each of the 19 locations indicated by the gold-colored circles. Each ΔV determination contributing to the contours is made as shown in Figure 5. The points at the base and peak of the velocity jump are determined by an automated analysis routine, and the midpoint of the rise is taken as the local CME front arrival time. The associated density and temperature jumps (Δn and ΔT) are determined and stored as well. (For uniform backgrounds this process is straightforward, but in the case of significant stream structure we subtract point by point from the CME run the corresponding interpolated values from a purely ambient simulation, run separately.) Note especially that the marker points denote the peak parameter jumps across the front, which is not typically the same as the shock jump. (In these coarse-grid simulations shocks are typically quite smeared, particularly where they are weak and/or oblique.) These data are collected for each sampling location in Figure 4, along with other metadata from each run, to be used in further processing.

Transit times for two members of an ensemble, computed as above, can also be represented by contour maps, as shown in Figure 6. Here we view the relative transit times at 1 AU across the front of the CME aimed directly at Earth (top) and another CME aimed 10°N and 10°W (bottom). Blue/red denotes later/earlier arrival time relative to that along the Sun-Earth line. The figure shows that because this CME is aimed away from Earth, its front crosses the 1 AU sphere at 10°N and 10°W a few hours before it arrives at Earth.

Collecting these data for all ensemble members into CSV-format files facilitates the creation of the transit time parameter plots shown in Figure 7. Since we will encounter many of these plots throughout the paper, we take pains here to explain what these plots are intended to convey. Figure 7 presents a 21-member subset of the full 81-member ensemble described above (we explain why we use a reduced, 21-member subset in section 5, below). Each symbol indicates the relation between the transit time of one of the CME ensemble members and the associated velocity (top) and density (bottom) jumps across that front at 1 AU along a radial line from the Sun to Earth. The colors indicate the launch speed, the round symbols indicate ensemble members aimed directly at Earth, and the squares indicate CMEs launched off the Sun-Earth line (here in varying combinations of 10° offsets in N and/or S pointing, not differentiated in this figure and intended only to suggest the resulting scatter in arrival time and strength). For both circles and squares, the size indicates the angular half width, according to the black scale located at top center on the figure.

We see that the slower CMEs (light olive green) take longest to get to 1 AU and produce relatively weak flow parameter jumps, whereas the fastest, biggest CMEs (dark green) arrive most quickly and produce sizeable jumps. Moderate-speed CMEs (green) fall in between, depending upon size and pointing. While the trends of the individual velocity groups are quite similar, there is a noticeable tendency toward lesser travel time for the faster ensemble members (more obvious in the density plot). Note that in this simple case of a uniform
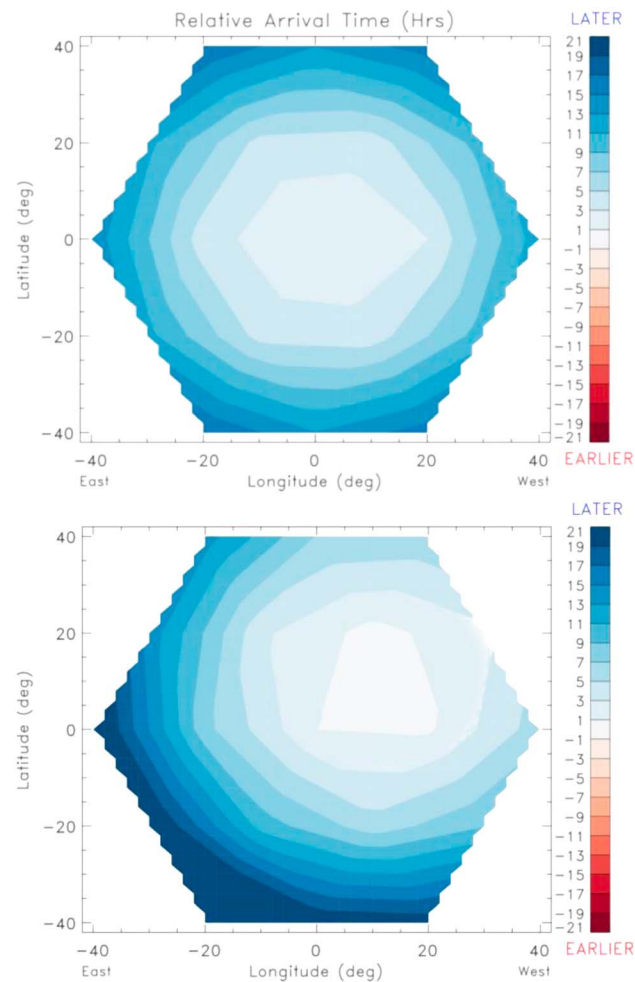
Figure 6. Latitude-longitude maps of relative arrival time at 1 AU for two model CMEs (for both, $V_{CME} = 800$ km/s, $\omega_{1/2} = 35°$) (top) one aimed directly at Earth, (bottom) the other 10° to the north and west. In both cases the blue contours indicate arrival times delayed relative to that at Earth.

HD background, all the off-pointed CMEs for each particular speed take longer to reach Earth, relative to those aimed directly at it.

Yet another way to gauge the 1 AU impact of the CMEs in our ensemble is shown in Figure 8, where—for those CMEs in the ensemble aimed directly at Earth—we view transit times (top), velocity jumps (middle), and density jumps (bottom) as a function of the angle from the CME center-line, stated in terms of the half width. Note that the angular widths of the velocity and density jumps at 1 AU are commensurate with the input half widths, with the amplitudes within $\pm \omega_{1/2}$ of centerline varying only slightly but dropping off rapidly beyond this range. This effect also becomes more accentuated with CME strength, whereas weaker CMEs show more rounded parameter jump profiles. The transit time plots (top), however, display decidedly broader, smoother profiles, even for the most powerful CMEs in this group. This difference is attributable to the fact that the transit time reflects directly the propagation of the shock front at the nose of the CME, whereas the velocity and density jumps—as measured here—include the effects of the ejecta, i.e., compressed ambient plus CME driver gas. This is consistent with what has been known for over 40 years [e.g., *DeYoung and Hundhausen*, 1973] that the shock front leading CMEs—being a wave—is intrinsically broader in angular extent than the ejected material driving it, much as the bow wave of a ship is wider than the vessel driving it. Thus, while CMEs off-pointed more than $\omega_{1/2}$ from Earth may still drive some disturbances at Earth, they mainly constitute "glancing blow" events or miss Earth altogether. In a practical sense, those CMEs whose estimated centerline offset from Earth (as given by $\psi_{CME} = \mathrm{acos}(\cos(\lambda_{CME})\cos(\varphi_{CME}))$) lies near and beyond $\omega_{1/2}$ should prove the most troublesome to predict accurately, since the impacts vary substantially in that range.

4. Ensemble Runs

In this section we present results of two extensive studies of hypothetical CME ensembles in unstructured and structured backgrounds covering a broad range of inputs, including variations in the ambient speed.

4.1. CME Ensembles in Uniform Hydrodynamic Ambients

We expand upon the demonstration ensemble introduced above to cover a broad range of CME parameter space, such as typically encountered in the course of operations. Figure 9 conveys a sense of the relevant parameter space. On the horizontal axis we have uniform ambients at three different levels (250, 350, and 450 km/s at R_0), while on the vertical axis we sort according to relative CME strength, which actually involves an amalgam of input speed (running from weak 300 km/s CMEs to very "strong" ones at 2000 km/s) and half width (and also pointing, relative to the 1 AU location of interest—so the real domain
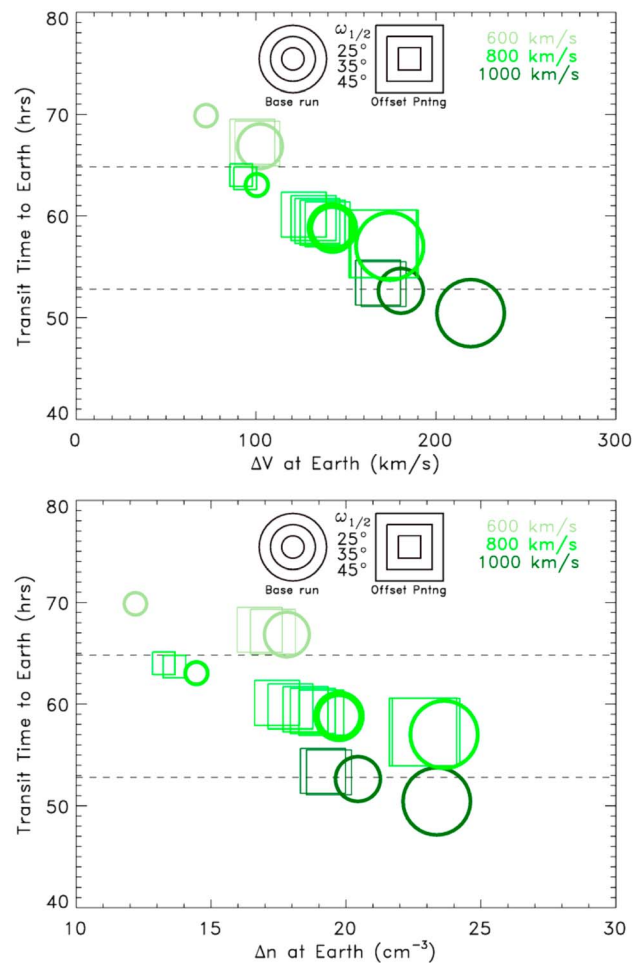
Figure 7. Scatterplots of arrival time at 1 AU versus (top) velocity jump and (bottom) density jump across the CME front for a 21-member ensemble launched into a uniform HD ambient with a moderate 350 km/s flow speed. Symbols in {light olive green, green, dark green} denote parameters for CMEs launched at {600, 800, 1000 km/s} with varying angular half widths $\omega_{1/2}$ and pointings relative to the Sun-Earth line: circles indicate those directed head-on, while squares indicate those at ±10° offset (a value chosen for illustrative purposes only). The size of the circles and squares indicates $\omega_{1/2}$ in each case. The plots show a well-ordered systematic variation from the weakest to strongest CMEs in the ensemble. (The motivation for using 21-member ensembles is detailed in section 5.)

is indeed a multispace that cannot be adequately represented here). Thus, in the lower left corner we have ensembles of slow CMEs injected into slow, dense ambient, while at upper right we have very fast CMEs propagating in faster, more tenuous ambient. Real CMEs and ambients may indeed fall outside these limits, but this range suffices to lay out the key issues in CME propagation (at least within the modeling realm).

For the purposes of this paper, we will focus upon the band of possible ensembles represented by the vertical red arrow (CMEs of varying strength propagating in a 350 km/s ambient) and by the green arrow (moderate-strength CMEs propagating in a range of ambients). Considering first ensembles in the vertical red arrow domain, the results in terms of the transit time versus parameter jump measure are presented in Figure 10. Here three sets of 21-member ensembles are depicted on a single plot, a blue-shaded set for slower, weaker CMEs, a similar green-shaded set for the moderate CMEs of Figure 7, and a reddish set for strong, fast CMEs. A fairly regular progression in the transit time versus parameter jump relation is seen in the velocity jump (Figure 10a), but something odd happens in the density jump plot (Figure 10b), where the density jump appears to approach some kind of limit (Δn is capped in the low twenties). Clustering about a limit like this can be a sign of chaotic behavior, but the most likely explanation is that shock heating is coming into play, with the stronger CMEs driving ever larger nonadiabatic temperature jumps at their leading edge. Figure 10d vividly confirms

that the gas pressure jump (ΔP_{gas}) across the CME front—the product of the density and temperature (ΔT) jumps (Figure 10c)—varies regularly with CME strength. Hence, the limit in the density jumps in Figure 10b stems from dissipative energy deposition across the CME shock front. The coarse grid probably overemphasizes heating in the shock compressions, which in reality should be narrower spatially and with a somewhat higher density. The Δn shown here should thus be taken as a lower bound and the ΔT as an upper bound.

We now turn to the role the ambient speed plays in the evolution of CMEs. Figure 11 displays an overlay of three ensembles, as in Figure 10, but this time all three feature the same moderate-strength CMEs (600/800/1000 km/s) launched into a different ambient (250/350/450 km/s at R_0). In Figure 11 (top), the run of bluish symbols near the top indicates the 600/800/1000 km/s set of 21 CMEs in a 250 km/s ambient, the next lower set (greenish) is the same set of CMEs in a 350 km/s ambient (again, as in Figure 7), and the lowest, reddish set is for the same CMEs in a 450 km/s ambient. Figure 11 (bottom) portrays the density
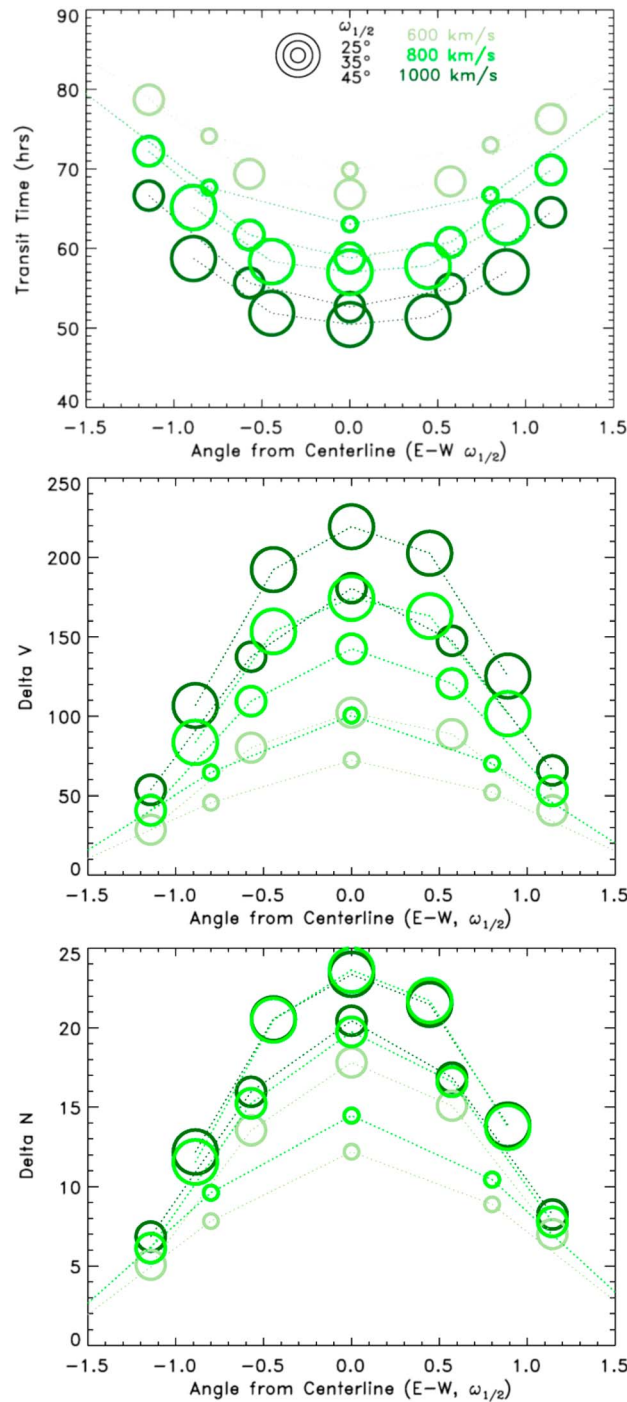
Figure 8. Angular distribution of (top) transit time, (middle) velocity jump, and (bottom) density jump relative to CME centerline at 1 AU for ensemble of Figure 7. The ordinate is presented in terms of CME half width (to provide broader context across CME). Color scheme denotes velocity at launch (as in Figure 7), while relative size of circle denotes $\omega_{1/2}$. Velocity and density jumps fall off sharply beyond $\omega_{1/2}$ from the CME centerline (reflecting the widths of the ejecta), whereas the transit time has a broader profile (since that is indicative of the shock width at the CME front).

amplitude at Earth for the three sets of data, in the same color scheme. The faster speed ambients are also less dense (see Table 1), so the pileup of material at the CME front is correspondingly reduced, whereas dissipative heating at the shock fronts rises (not shown). Comparing Figures 10 and 11, we see that the speed of the ambient background can have nearly as much to do with transit time and 1 AU CME front properties as intrinsic CME properties (launch velocity, half angle).

4.2. CME Ensembles in Structured Ambients

Accounting for the interaction of CMEs with preexisting stream structure is a basic element of space weather forecasting. Just where a CME encounters a dense stream front with its associated velocity gradients obviously affects the propagation characteristics, and this sort of interaction is routinely taken into account in operational CME forecast modeling efforts at SWPC and by other groups around the world. Although some exploratory studies in this direction have been published [e.g., *Case et al.*, 2008; *Lee et al.*, 2013, 2015], no comprehensive modeling effort has been attempted. In particular, what has been lacking is an assessment of 1 AU CME properties in the spirit of that undertaken in section 4.1 but which includes interaction with significant, representative stream structure. It is the intent of this section to lay the basis for the interpretation of ensemble modeling of CME propagation in realistic ambient flows that include significant stream structure.

4.2.1. CME Ensembles in a Tilted-Dipole Stream Structure

The most useful generic background stream structure for this purpose is the tilted-dipole configuration described in *Pizzo* [1982, 1991]. In view of the operationally oriented nature of this study, we will simplify the specification of the input stream structure at $R_0 = 0.1$ AU as follows. First, we define the velocity structure in a coordinate system that assumes no tilt of the stream configuration to the solar equator (i.e., velocity and density are functions of heliolatitude λ only) as

$$V = (V_{\text{fast}} - V_{\text{slow}}) \times \sin^p(\lambda) + V_{\text{slow}},$$
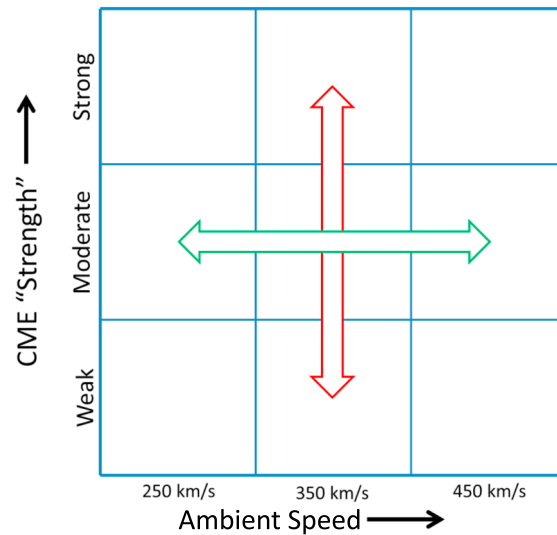
Figure 9. Schematic of CME input parameter space to be explored in section 4.1. The term "strength" refers to a combination of input CME speed and half width (which, in turn, is correlated with mass). Along with the CME pointing relative to the observer, the possible inputs thus actually constitute a hypercube, which for simplicity has here been collapsed to a square. This part of the study thus covers everything from "weak" CMEs in a slow ambient (lower left corner) to "strong" CMEs launched into a "fast" ambient (upper right corner). For tractability, we present here only results indicated by the red (range of CME strength in moderate-speed ambient) and green (moderate-strength CMEs over range of ambients) arrows.

where p is an arbitrary constant used to sharpen the velocity distribution about the equator and is set here to 16, $V_{fast} = 600$ km/s, and $V_{slow} = 250$ km/s. The density is given by

$$n = n_{fast}\left(\frac{V_{fast}}{V}\right)^2,$$

and the total (gas only) pressure is held constant on R_0, with the result that the temperature variation is the inverse of the density variation. Given these specifications, it is trivial to impose any desired tilt of the coordinate system (and thus the large-scale flow parameter configuration) via the appropriate 3-D coordinate rotation.

The result, for a 30° tilt, is the velocity and density distributions on R_0 shown in Figure 12. Here we view a sinuous band of high-density, low-speed material weaving about the solar equator. Quantitative results quoted in this paper clearly depend upon the specified tilt and would also vary if some other geometry were imposed. However, configurations like this typically apply over a substantial part of the solar cycle, and imposing this assumption also facilitates interpretation via comparison with papers already in the literature.

The first simulation at hand is indicated by the red circle in the middle of both plots. This represents the location and cross section of CME material (parameterized as in section 4.1) that is injected into the structured tilted-dipole ambient at some time T_0 after the full 3-D ambient numerical solution has evolved into a stable, fixed state. Just before CME launch, the structured 3-D ambient is characterized by two large, steady-stream structures, one (near 0° longitude) with a northward leading tilt (i.e., the normal to the front points up and to the right, in the sense of rotation) and the other (at ±180°) with a complementary southward leading tilt. Thus, the properties of the steady, evolved, pre-CME ambient are similar to that discussed at length in *Pizzo* [1982, 1991].

This case, where the CME is injected directly into the heavy stream material at the center of the plot, ensures the maximum interaction of the CME with the stream. A major portion of the CME will plow through the densest part of the stream right near the inner boundary, thereby experiencing the maximum slowing by momentum exchange. As the CME propagates out from R_0, the residual extra momentum in the CME material will break out around the dense stream, as parts of the CME above and below the initial center of the CME will be less encumbered by the dense stream front material and race out ahead of the slower part near CME center. Eventually, the injected CME material flattens into a distorted, bent shape like that depicted in Figure 2 of *Odstrcil* [2009].

The flow pattern at 1 AU is illustrated in Figure 13 (top), which shows a ΔV plot like Figure 7 for the same CME input speed and $\omega_{1/2}$ as that case but launched directly into the midst of the stream structure depicted in Figure 12. The ΔV distribution at 1 AU is bifurcated as a consequence of interaction with the dense, heavy stream component lying across its path near the Sun. The dark, low-ΔV region running diagonally from upper left to lower right marks the impeded flow in and around the dense stream front material, whereas the bright contours of high ΔV indicate areas off to the side (upper right, lower left) where the interaction between the stream front and the CME is much weaker. The associated arrival time plot (Figure 13, bottom) is color scaled to show the delayed (blue) arrival along the low-ΔV corridor; the earliest (red) arrivals for this CME are coincident with the high-ΔV band running from lower left to upper right. More generally, the precise distributions
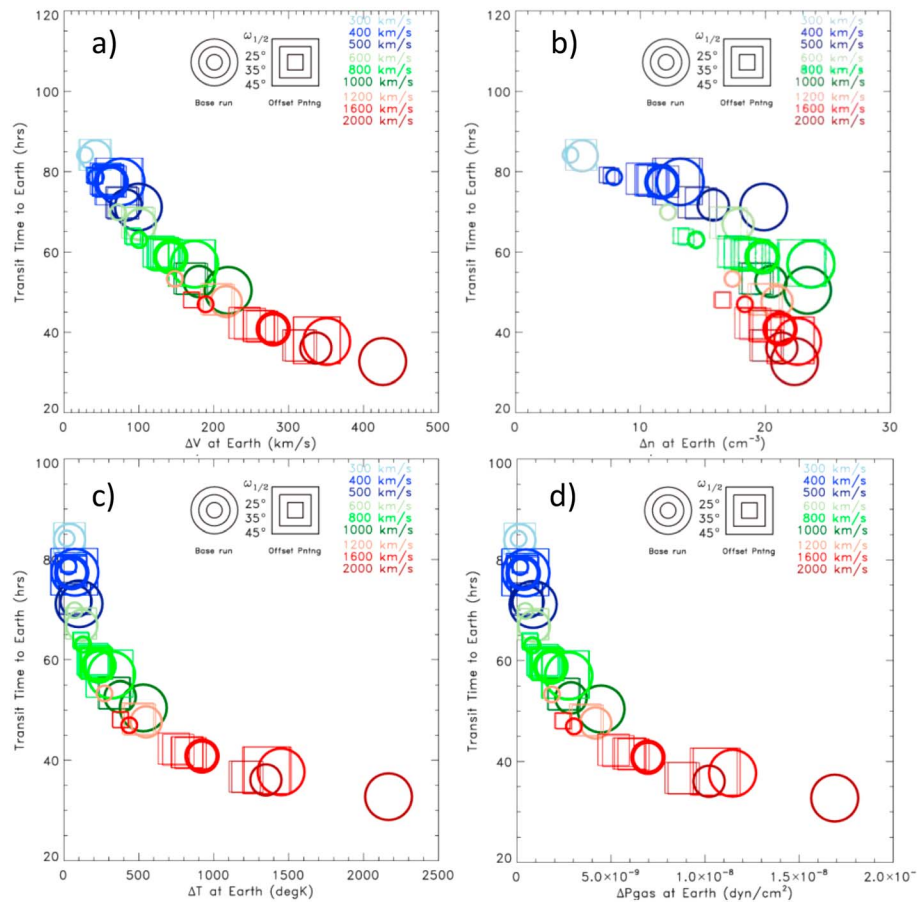
Figure 10. Color-coded scatterplots of transit time versus fluid parameter jumps for three 21-member ensembles covering input CME speeds between 300 and 2000 km/s, injected into a 350 km/s ambient. (a) ΔV, (b) Δn, (c) single-fluid temperature jump (ΔT), and (d) gas pressure jump (ΔP_{gas}). The moderate-speed CME data are in green (as in Figure 7), while blue signifies a slower ensemble and red a faster ensemble, as indicated at upper right. Square and round symbols are defined as before. The 1 AU velocity jump increases monotonically with input speed, while the number density jump reaches an apparent limit of ~25 cm^{-3}. The continuity in ΔP_{gas} shows that the limit on Δn stems from shock heating (as exemplified by the monotonic increase in ΔT) and is not indicative of the kind of clustering anticipated in a chaotic system.

of the transit time and ΔV will vary from CME to CME, depending upon the pointing relative to the dense stream material and the size and speed of the CME. In addition, the magnitude of the interactions will vary for weaker to stronger velocity and density contrasts and stream topologies near the Sun.

To assess quantitatively the 1 AU properties of CMEs propagating through the structured ambient of Figure 12, three 21-member ensemble sets are injected into it, with CMEs of varying size and speed being launched over an angular spread of ±10° about the center of the red circle in Figure 12. The result is seen in Figure 14, which is in the same format as Figure 10, for reference. Here we see the now-familiar arc of members in the TT-ΔV plot but a more sharply truncated TT-Δn distribution relative to that in Figures 10a and 10b. Again, shock heating, as measured by ΔT and ΔP (not shown but similar to Figures 10c and 10d) is at play in imposing the peak Δn limit. The ensemble distribution in Figure 14 is very regular and systematic, with no clumping or other indications of chaotic behavior.

4.2.2. Ensemble Interactions for CMEs Injected at Different Locations in the Stream Structure
The 21-member ensemble in section 4.2.1 is launched into the midst of the dense stream material where it crosses the equator, guaranteeing the strongest possible interaction with the background structure. We now consider the more general case of ensembles launched at various offsets about the centroid of that ensemble.

First, we address the case of ensembles injected at small offsets (±20° in both latitude and longitude) about the stream front. The results of this experiment provide a feel for the sensitivity of the previous results to CME directionality in the presence of a strong background structure. We find that scatterplots in ΔV and Δn (not shown)
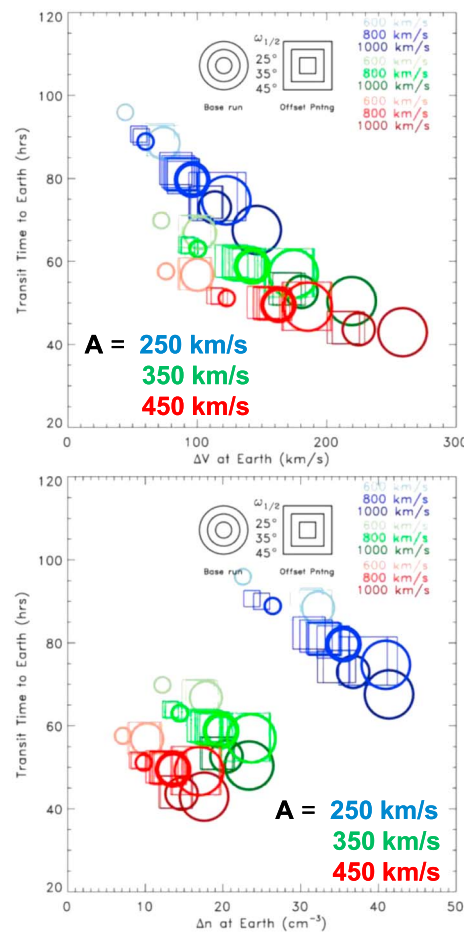
Figure 11. Scatterplots of (top) ΔV and (bottom) Δn for three 21-member ensembles injected at moderate speeds into ambient flows of {250, 350, 450 km/s}, which are color coded {blue, green, red}. The slowest ambients have higher intrinsic densities, while the faster ambients are progressively more tenuous (see Table 1). Hence, a fast CME in a slow ambient will get bogged down as it sweeps up considerable material at its front, while the same CME injected into a faster, more tenuous ambient will accumulate less material and transit more rapidly to 1 AU. On the other hand, the ambient also acts somewhat as a conveyer belt—a slow CME launched into faster ambient wind will propagate to Earth more quickly than its intrinsic speed might suggest.

are quite similar in range and form for all six of the offset ensemble members, looking much like that in Figure 14. There is slightly more differentiation according to speed and size in those ensembles injected just ahead of the stream front and rather less for those injected behind, but the effects are overall very minor. More scatter occurs in the Δn than in the ΔV plots for these offset pointing cases, but all exhibit approximately the same "knee" in the Δn distribution as in Figure 14 (bottom).

But what of CMEs that are injected farther from the slow stream front? The general run of results can be understood by consideration of sample CMEs introduced at various places with respect to the CIR front. Figure 15 presents plots of ΔV and Δn for a CME configuration {$V_{cme} = 1000$ km/s, $\omega_{1/2} = 35°$} injected in the equatorial plane at five discrete locations [$\varphi = \{80°, 40°, 0°, -40°, -80°\}$] relative to the nominal Figure 12 stream front, as marked by the colored crosses in that figure. In each case, the transit times and jump parameters every 20° across the front are denoted by the colored symbols. From top to bottom, the plots represent the transit time TT to 1 AU, the local velocity jump ΔV, and the density jump Δn. The black curves near center of each plot indicate 1 AU CME properties for the case where the CME is injected directly along the CIR front near the Sun (as in Figure 14). There, the density jumps are relatively high (since the CME is running directly into the stream front), while the velocity

jumps are muted; the somewhat asymmetric peaks in both are localized about the CIR front. The same CME injected far from the CIR front ($-80°$, purple; $+80°$, blue) propagates into a broad area of fast stream and therefore evolves nearly as it would in a uniform, fast background.

For a CME injected 40° ahead or behind the CIR front, however, the story is more complicated. For a CME injected 40° to the west (red, positive longitudes) of the CIR front, the easternmost portion ($\varphi \approx 0°$) of the CME is greatly weakened (very low ΔV) and slowed through strong interaction with the CIR, which also pushes up the already high CIR densities. On the other hand, the western segments of this CME expand freely into the back of the preceding fast stream, such that the peak ΔV location occurs some 20° to the west of that CME centerline. Conversely, the CME launched 40° to the east of the CIR front (green, negative longitudes) runs up upon the CIR from behind and exhibits a basically inverse evolution. That is, the peak ΔV location now occurs 20° to the east of the CME centerline, while high Δn is experienced in the vicinity of the CIR. Similar plots can be constructed for faster or slower and narrower or wider examples, with corresponding variation in the angular response.

Finally, in terms of transit time versus ΔV and Δn plots, the most interesting deviations from Figure 14 (CMEs in the vicinity of the stream front) and Figure 10 (CMEs in uniform flow far from a stream front) occur near
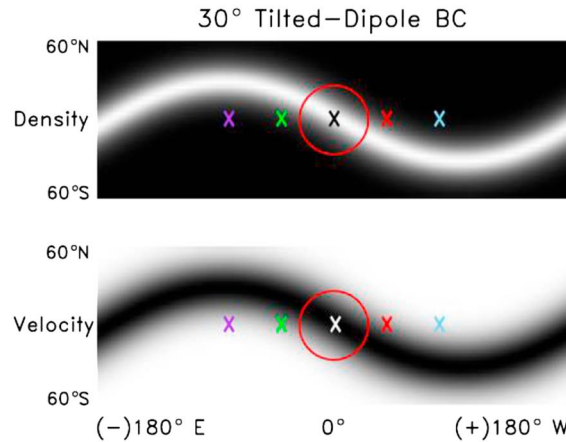
Figure 12. Tilted-dipole ambient flow geometry used to illustrate how an idealized, highly structured solar wind background may affect CME propagation. This kind of structure is most commonly associated with the declining phase of the solar cycle. The centroid of the ensemble discussed in section 4.2.1 is indicated by the black/white cross in the red circle; the other colored crosses denote the centroids of similar, additional ensembles discussed in conjunction with Figure 15.

±40° relative longitude. Figures 16a and 16b illustrate how ensemble sets of moderate and fast CMEs react to being launched at 40° W into the rarefaction region trailing the fast stream preceding (in the sense of rotation) the CIR front; conversely, Figures 16c and 16d show the case where the CMEs are launched into the fast flow 40°E, behind the CIR front. In the former, it can be seen that the CME fronts develop very large ΔV but also very small Δn, since those CMEs are accelerating "downhill" into the very tenuous material at the back of the preceding fast stream. In the latter, however, where the CMEs are overtaking the stream front compression, we see that the Δns become considerably larger even as the ΔVs are somewhat reduced, though still substantial.

The net of all this is that in the presence of a sufficiently strong CIR background, the 1 AU manifestation of any given CME can depend greatly upon where it is injected with respect to that particular CIR structure. The severity of the interaction also depends upon the input strength of the CME (as given by its speed and half angle) relative to the velocity-density contrast of the ambient structure. That is, for cases (not shown) where the amplitude of the CIR structure is reduced (by decreasing the velocity and/or density contrast in Figure 12), the 1 AU impact of any given CME will vary, such that the structure in the ambient at some point has little influence on the CME propagation.

5. Taylor Analysis

To complement the expositions of transit time versus parameter jump plots presented above, we develop a mathematics-based way of describing and assessing the results (as opposed to a graphical approach, e.g., as in *Taktakishvili et al.* [2010]). Since we are dealing with coarse-grid solutions that are inherently smooth and regular, it makes sense to adopt a Taylor expansion approach for the analysis of our numerical solutions. In this way, we can quantify systematically the relative contributions of the various input parameters individually and in concert, and we can distil the analysis of any given ensemble to one simple graphic.

Viewing the CME transit time as a function of the four input cone parameters leads naturally to the following second-order Taylor expansion for dT, the (signed) variation in the transit time to 1 AU directly attributable to changes in the inputs:

$$\mathrm{d}T = \frac{\partial T}{\partial v}\delta v + \frac{\partial T}{\partial \omega}\delta\omega + \frac{\partial T}{\partial \lambda}\delta\lambda + \frac{\partial T}{\partial \varphi}\delta\varphi$$

$$+ \ 1/2\left[\frac{\partial^2 T}{\partial v^2}\delta v^2 + \frac{\partial^2 T}{\partial \omega^2}\delta\omega^2 + \frac{\partial^2 T}{\partial \lambda^2}\delta\lambda^2 + \frac{\partial^2 T}{\partial \varphi^2}\delta\varphi^2\right]$$

$$+ \ \left[\frac{\partial^2 T}{\partial v\partial\omega}\delta v\delta\omega + \frac{\partial^2 T}{\partial v\partial\lambda}\delta v\delta\lambda + \frac{\partial^2 T}{\partial v\partial\varphi}\delta v\delta\varphi\right]$$

$$+ \ \left[\frac{\partial^2 T}{\partial \omega\partial\lambda}\delta\omega\delta\lambda + \frac{\partial^2 T}{\partial \omega\partial\varphi}\delta\omega\delta\varphi + \frac{\partial^2 T}{\partial \lambda\partial\varphi}\delta\lambda\delta\varphi\right],$$

where all the derivatives like $\partial T/\partial v$ and $\partial T^2/\partial v^2$ are evaluated from centered finite differences derived from the solutions for the various ensemble members (e.g., $v_0, v_0 \pm \Delta v$) and $\delta v, \delta\omega$, etc., are observational error estimates, as described in the next paragraph. From the associated finite difference expressions it can quickly be ascertained that only 21 of the full set of 81 runs are needed to second order (see Table 2 and caption),
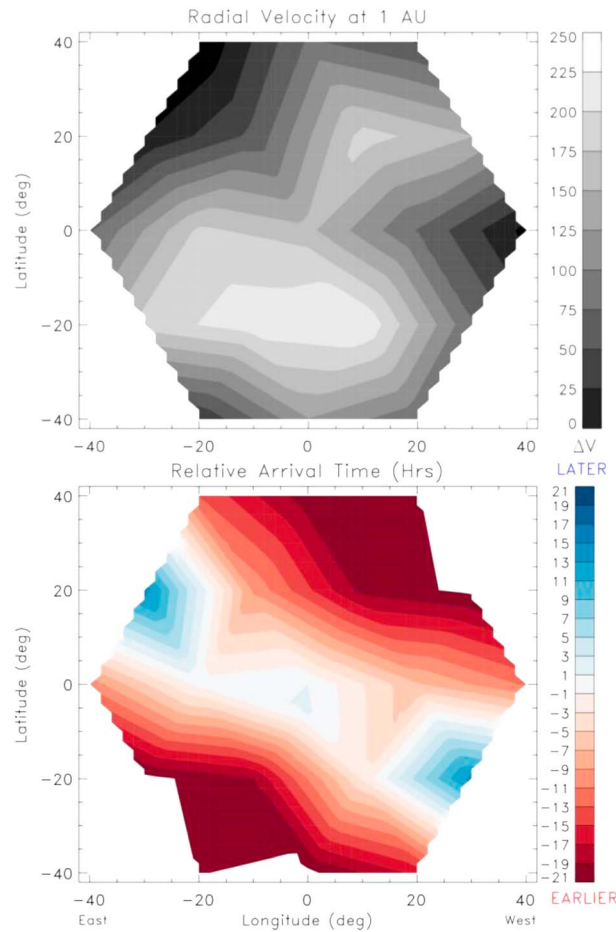
Figure 13. Latitude-longitude maps of (top) ΔV and (bottom) 1 AU arrival time for a {$V_{CME} = 800$ km/s, $\omega_{1/2} = 35°$} CME launched into the ambient at the center of the red circle in Figure 12. The presence of the structured ambient profoundly affects the CME properties near Earth, as compared to the same CME launched into a uniform ambient (cf. Figure 6)—where as the CME encounters the CIR front it is slowed down considerably, while away from the front it propagates much faster.

representing a considerable reduction in simulation time. (Because of the coarse grid used in the simulations, little to nothing is lost in the reduced set.)

It is essential to understand that the derivatives in the above expression are evaluated from the model runs (based upon the given model spreads Δv, etc.), but the net dT requires accurate estimation of the error in the input cone parameters (δv, etc.). These errors come independently from observations and experience with the cone-fitting method used; for any given CME we also need to know how well we think we have estimated the velocity, the half width, and the direction. These depend upon many factors for CMEs, ranging from isolated weak disturbances to highly energetic full-halo events, where the disturbed background corona can make their estimation very difficult. Unavailability of NRT multiview spacecraft images or poor coronagraph configurations (such as at present, when the STEREO spacecraft lie nearly along the Sun-Earth line) also contribute to the error estimates.

Figure 17 depicts the results of the Taylor analysis at three different locations relative to CME centerline, for the simple ensemble of CMEs injected into a uniform hydrodynamic background (Figure 7). From the left in Figure 17, each set of colored symbols presents, in order, the (signed) magnitude of the transit time uncertainty stemming from the four first derivatives (symbolized as {v, ω, λ, φ}), then the four second deriva-

tives {vv, $\omega\omega$, $\lambda\lambda$, $\varphi\varphi$}, and finally the six cross derivatives {$v\omega$, $v\lambda$, $v\varphi$, $\omega\lambda$, $\omega\varphi$, $\lambda\varphi$}. These are computed from $(\partial T/\partial v) \, \delta v$, and so on, for each term in the Taylor expansion. In the figure, "Mdl ΔP" is the ensemble parameter spread {Δv, $\Delta\omega$, $\Delta\lambda$, $\Delta\varphi$}, and "Est δP" is the estimated magnitude of the uncertainty in each parameter, not necessarily the same as the ensemble model spread. (Here we simply posit a reasonable guess at Est δP, as indicated in the figure, for purposes of exposition only.) Finally, "RMS$_1$" refers to the net root-mean-square error for the first-order terms only, since that is where the bulk of the uncertainties come from.

In Figure 17, green denotes the 14 Taylor components along the CME centerline (i.e., $\varphi = 0°$), brick red the components evaluated along $\varphi = +20°$ to the west, and purple along $\varphi = +40°$. (Evaluation to the east is essentially the same but with the φ component contributing negative dT.) Discussing first the Taylor components along the centerline of this ensemble of moderate-strength CMEs (green symbols), the largest contribution of uncertainty comes from the first-order velocity term, then the half width, and lastly the second-order velocity and half width (of opposite sign). All the other contributions are quite small. We conclude that for the range of parameters applying to this ensemble, the uncertainties in the CME velocity and half width dominate the accuracy of any forecast. Pointing is a much lesser issue, at least within the range of offsets considered. Taylor analysis for the same CMEs viewed from +20° (brick red) and +40° (purple) to the west (or, equivalently, for a CME offset by the same amount to the east), shows rapid increase in the φ and ω contributions. On the basis of Figure 8, this should come as little surprise, as the uncertainties
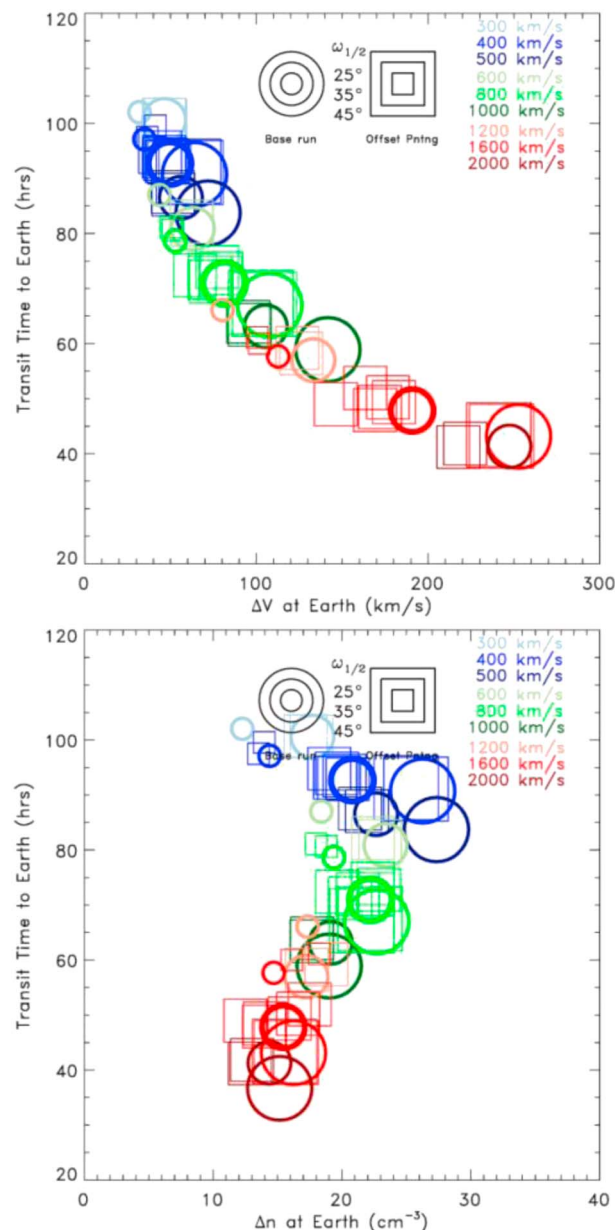
Figure 14. Scatterplots of arrival time at 1 AU versus (top) velocity jump and (bottom) density jump across the CME front for three 21-member ensembles launched into the middle (center of red circle) of the structured ambient of Figure 12. Compared with Figure 10 (the uniform ambient case), it is evident that the CMEs propagate a bit slower to 1 AU and the velocity jumps are reduced, all by virtue of interaction with the dense material at the core of the structured ambient front. The peak density jumps in the slower (blue) CMEs are somewhat increased, but those in the faster sets (green and red) reflect the onset of serious shock heating and are systematically reduced with increasing CME speed.

increase dramatically for glancing blow CMEs seen farther off centerline. (Were these CMEs displaced from the Sun-Earth line in latitude, the Taylor λ component could likewise become significant.)

The results from such analyses can vary from ensemble to ensemble, with the magnitude and sign of the v and ω derivatives and those of λ and φ switching about, depending upon the specific conditions of the ambient stream structure. It is the overall picture of the variation in sensitivities that matters, as will now be developed.

Figure 18 (top) depicts the relative magnitude of the Taylor components (henceforth we evaluate only along the ensemble axis) for ensembles of slow (300–500 km/s; blue) CMEs, moderate-speed CMEs (600–1000 km/s; green, from Figure 17), and fast (1200–2000 km/s; red) CMEs propagating into a uniform 350 km/s background. We see in Figure 18 (top) a systematic progression from slow CMEs, where the input speed dominates the possible variation in transit times, to fast CMEs, where the CME half width becomes more important. In the former case, the interaction plays out gently over a longer propagation time, so the relative difference between CME and background speed tends to be preserved; in the latter, the interaction between CME and ambient is so strong and impulsive that the relative momentum, as set by the input half width, determines how fast the CME can punch through the background flow.

To assess in the same way the impact that differences in the ambient speed have upon these processes, we compare in Figure 18 (bottom) the Taylor component plots for moderate-speed CMEs (600–1000 km/s) propagating through slow (250 km/s; light blue), moderate (350 km/s; green), and fast (450 km/s; gold) uniform backgrounds. Here we find that in the slow ambient (light blue) the dominant inputs are the CME speed and width, yet for the fastest ambient (gold) it is mainly

the CME speed. Physically, what is happening is that since the density in the slow ambient is roughly thrice that in the fast ambient, the ability of a CME to punch its way through the dense slow wind depends upon the input CME mass and momentum (V and $\omega_{1/2}$), whereas the lightweight background in the fast ambient poses much less of an obstacle, enabling the CME to barrel along relatively unimpeded by the solar wind ambient. It should also be noted that the RMS$_1$ uncertainties in transit time
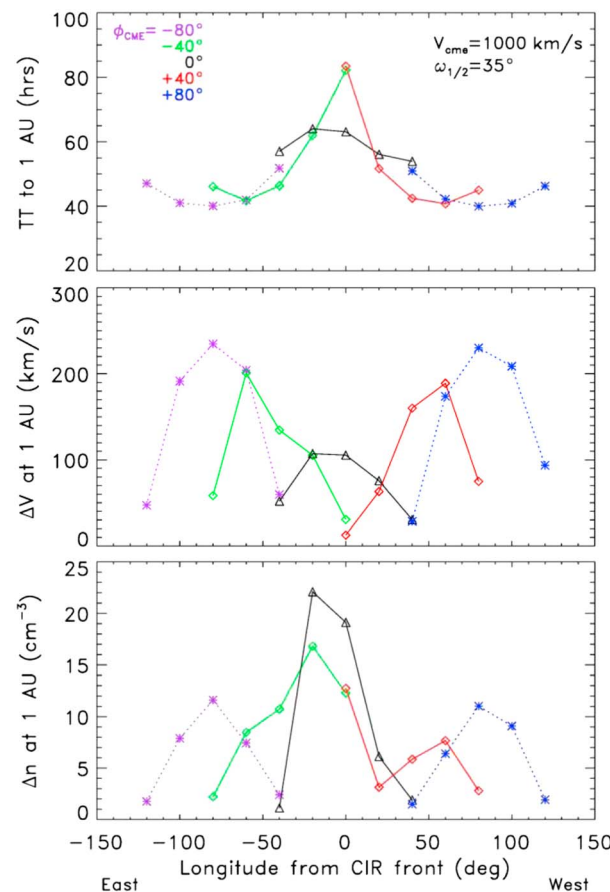
Figure 15. Illustration of the spatial variation of CME front properties spanning a variety of ambient structures. For each set of colored curves, the symbols mark (top) the transit time, (middle) the velocity jump, and (bottom) density jump in the equatorial plane at the indicated longitudes with respect to the CME front. For example, the black curves depict CME front properties for the {V_{CME} = 1000 km/s, $\omega_{1/2}$ = 35°} case where the CME is launched into the midst of the ambient CIR front. Although the transit times across the front do not vary too much, the velocity and density jumps do. For the same CME launched ±80° from the CIR front (purple; blue), where it propagates freely into more uniform ambient conditions, more regular, rounded variations are seen. The same CME launched at intermediate distances from the CIR (green; red) show much more distortion due to variations in the localized interactions.

attributable to variations in the assumed ambient speed are (at least over the range 250–450 km/s) not substantially different from those due to input speed and mass (Figure 18, top).

Taylor analysis for the tilted-dipole ambient configurations we have considered (section 4.2 and Figures 13–15) reveals that for modest offsets about a substantial stream front near the Sun (i.e., ensembles centered at the yellow locations), a high degree of order nevertheless persists in the distribution of the various Taylor components. The primary results are illustrated in Figure 19 (top), which compares the Taylor analysis for the ensemble of moderate-speed CMEs (top) centered at the heart of the red circle in Figure 12 to that of a fast ensemble (Figure 19, bottom) similarly centered (i.e., the moderate and fast ensembles depicted in Figure 14). Each plot presents Taylor components assessed at 1 AU along the equatorial plane at five different longitudes (0°, ±20°, and ±40°) with respect to CME center. (Here black symbols represent Taylor analysis taken along 0°, with Taylor analyses along offset directions being color coded as indicated in the RMS$_1$ insets.) For the moderate-speed CMEs (Figure 19, top), the five subplots exhibit considerable and varied scatter, with the first-order components clearly dominating. For the fast CMEs (Figure 19, bottom), even more scatter is evident, and the half width and east-west pointing components become more prominent. The net effect is that in the presence of strong ambient structure, error in all the CME inputs can have an effect, with the relative proportions depending upon the observing position.

Taylor analyses for CMEs offset much farther from the CIR front (not shown) exhibit similar behavior, though differing somewhat in detail. From this we conclude that the utility of Taylor analysis will in general depend nearly as much upon accurate knowledge of the ambient structure as upon that of the CME inputs. Careful study and analysis of historical data sets will hopefully provide telling guidance in this regard.

Finally, we note that Taylor component plots for ΔV and Δn may be generated the same way to quantify the uncertainties in parameter responses other than transit time, but that would be best addressed against historical data.

6. Summary

We have established a general basis and methodology for evaluating systematic relations among projected 1 AU CME properties across a broad range of near-Sun inputs used in an operational forecast model. In particular, we have (a) justified why the modeling of CMEs for forecast purposes may be approached in the
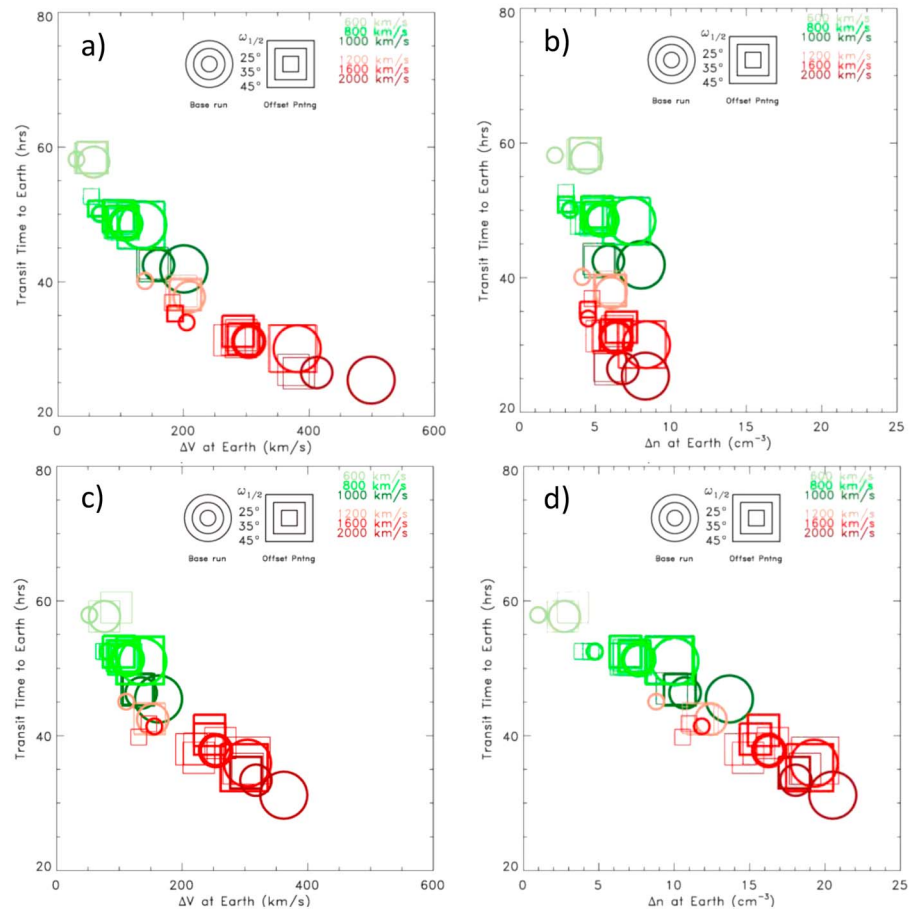
Figure 16. Medium (green) and fast (red) 21-member CME ensembles centered on the equatorial plane and (a and b) 40° to the west and (c and d) 40° to the east of the ambient CIR. Because the westward set can propagate out ahead of the CIR front into the back of the preceding fast stream, its velocity jumps are high while the density jumps are relatively low. The CMEs launched to the east, however, eventually encroach upon the rear of the CIR front, where they are slowed due to the pileup of material there.

coarse-grid HD limit; (b) included CMEs across a wide range of input parameters; (c) considered both uniform and structured background cases, in on-axis and off-axis regimes, over a range of presumed ambient flow speeds; (d) developed both graphical and analytic approaches for interpreting the results, for both research and operational applications; and (e) provided a mathematical basis for ensemble composition, including the specification of observational error estimates in the CME inputs.

Among the multitude of ensembles considered (including many not shown), we invariably found regular patterns in dT versus ΔV and dT versus Δn across a wide range of inputs for CMEs launched into various uniform and structured backgrounds. While monotonic behavior was the norm, a noted exception was the characteristic "bend" or limit in Δn as a function of dT (Figure 10), which we demonstrated was due to shock heating in the model and is definitely not clustering indicative chaotic behavior. These same trends persist in test runs where the spiral magnetic field was included (not shown), with only minor adjustments to Δn as the magnetic field compression in that case also contributes to the ΔP_{total} across the CME front.

We infer from all this that in practical terms the evolution of CME fronts to 1 AU appears to be a nonchaotic process, wherein the output is directly related to the input. That is, over even a large range of variations in inputs regular, repeatable, predictable variations in outputs accrue.

We find no indication of chaotic behavior (such as clustering of results), even when there is significant interaction with structured ambients. The main uncertainty in forecasting thus comes from uncertainty in inputs

Table 2. Specification of the Individual CME Runs Needed to Support Second-Order Taylor Analysis of the Ensemble Sensitivity to the Inputs[a]

Member #		v_0+dv	v_0	v_0-dv	$\omega_0+d\omega$	ω_0	$\omega_0-d\omega$	$\lambda_0+d\lambda$	λ_0	$\lambda_0-d\lambda$	$\phi_0+d\phi$	ϕ_0	$\phi_0-d\phi$
1	Base (and d2/dx2)		x			x			x			x	
2	d/dv	x				x			x			x	
3				x		x			x			x	
4	d/dω		x		x				x			x	
5			x				x		x			x	
6	d/dλ		x			x		x				x	
7			x			x				x		x	
8	d/dφ		x			x			x		x		
9			x			x			x				x
10	d2/dvdω	x			x				x			x	
11				x			x		x			x	
12	d2/dvdλ	x				x		x				x	
13				x		x				x		x	
14	d2/dvdφ	x				x			x		x		
15				x		x			x				x
16	d2/dωdλ		x		x			x				x	
17			x				x			x		x	
18	d2/dωdφ		x		x				x		x		
19			x				x		x				x
20	d2/dλdφ		x			x		x			x		
21			x			x				x			x

[a]Along with an ambient run, each ensemble member is a variant about a "base" CME run (subscript "0"). Only 21 CME runs are needed, since the second-order cross derivatives can be efficiently computed as $f_{x,y}(x,y) \approx \dfrac{f(x+h,y+k)-f(x+h,y)-f(x,y+k)+2f(x,y)-f(x-h,y)-f(x,y-k)+f(x-h,y-k)}{2hk}$ where $\{x, y\}$ represent any two of the physical variables $\{v, \omega, \lambda, \phi\}$, and $\{h, k\}$ are the corresponding $\{\Delta x, \Delta y\}$. (From http://en.wikipedia.org/wiki/Finite_difference.)
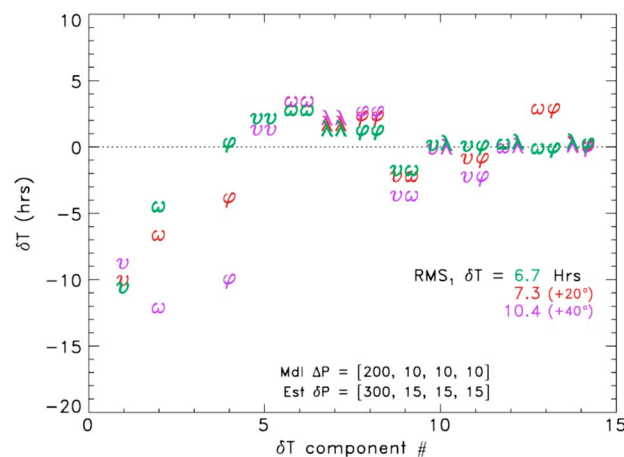

Figure 17. Taylor analysis diagram for a 21-member ensemble launched into a uniform, 350 km/s ambient (cf. Figure 7). From the left, each Greek letter corresponds to the value of the signed first derivative of each of the four parameters (v, ω, λ, φ), followed by their second and mixed second derivatives, as indicated by double symbols (e.g., vv, and $v\omega$). The green set of symbols represents the derivative values evaluated head-on to the CME; brick red and purple sets correspond to Taylor components evaluated 20° and 40° off axis, all in the equatorial plane. "RMS$_1$ δT" values refer to the root-mean-square of the first-order terms for each set, as indicated by color. "Mdl ΔP" is the corresponding model spread used in the ensemble (e.g., ±200 km/s, ±10° for the angular parameters), and "Est δP" is the estimate of the uncertainty in the input CME parameters (as supplied from empirical analysis of operational CME fits). In each case, net error in the velocity estimate remains about 10 h, while that in the half angle and east-west pointing becomes important off axis to the CME centerline (cf. Figure 8).

with respect to CME parameters *and* to the ambient state. On a physical basis, this outcome makes eminent sense, since we are dealing here with momentum-dominated flows in the hypersonic regime, wherein quite simple physics dominates the large-scale evolution with which we are concerned.

These findings constitute a basis for developing improvements in our forecasting abilities. Explicitly, the primary advances will come from better characterization of CME and ambient inputs, as opposed to better propagation models. They also suggest that some simplification in forecasting for very fast, high-mass CMEs may be possible, in that a lookup table approach may provide a viable forecast strategy, as for tsunamis. Such a strategy would be limited to the most energetic events, since the ambient should then not matter much—but these are the very cases that demand immediate, accurate forecasts. Given that the structured ambient considered in this study is near a worst case example, it must be left to the NRT experience to see if and when specific consideration of the ambient structure may be dispensed with and a
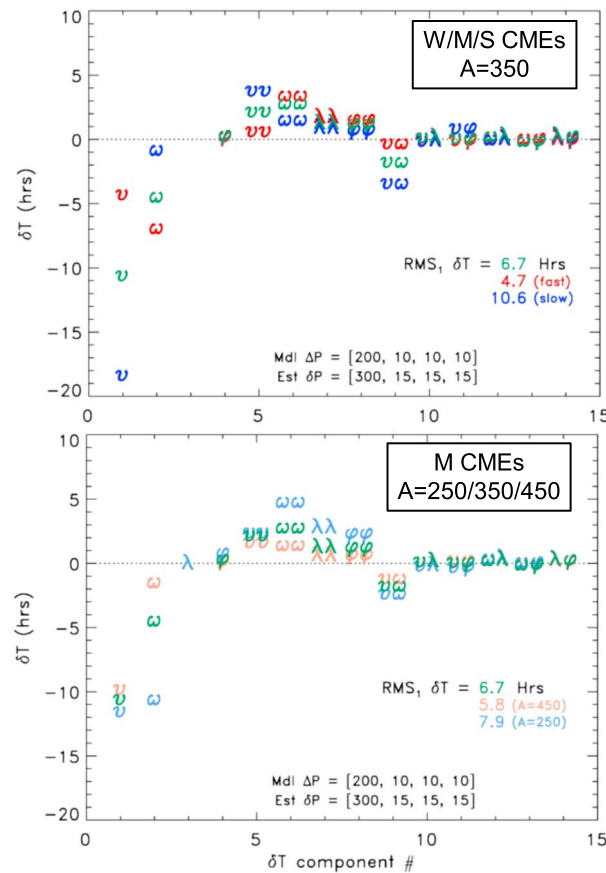
Figure 18. (top) A color-coded Taylor analysis diagram for weak (W, blue), moderate (M, green), and strong (S, red) CME 21-member ensembles launched into a uniform 350 km/s ambient. It can be seen that error in the CME velocity estimate dominates transit time error for slower CMEs. (bottom) Taylor analysis diagram for a moderate 21-member ensemble launched into uniform $A = \{250, 350, 450$ km/s$\}$ ambients. Green corresponds to the on-axis 350 km/s ambient of Figures 10 and 11, while gold and light blue depict Taylor components for CMEs launched into 450 km/s and 250 km/s ambients, respectively. While the error in transit time attributable to error in the CME velocity estimate does not vary much (assuming the ambient speed is known accurately), uncertainty in angular size becomes as important for slow ambients.

representative uniform outflow may suffice. For the immediate future, it would be prudent to run coarse-grid ensembles routinely to derive the likely spread in outcome for any given CME.

To that end, one of the more important outcomes of this study is the development of reliable statistical means to evaluate the anticipated dT for any given ensemble. We introduced a Taylor expansion approach to enable the identification of the input parameter or parameters exerting the greatest leverage on the transit time forecast. Moreover, while the ensemble pointing spreads $\Delta\lambda$ and $\Delta\varphi$ and the observational uncertainties δ employed here were for illustrative purposes only, we stressed that the formal error associated with each input parameter must be specified from observational experience. It can only be expected that these error estimates may vary widely. Where a CME has a well-defined shape and there are sufficient images from multiple well-situated spacecraft, we can expect the error to be minimal. For truly impulsive, fast events with a strong halo component, having only two views with spotty image availability, the error can expand substantially. These debilitations are a fact of life in NRT forecasting, so any feasible approach must incorporate allowance for error in inputs—adjustable on a case-by-case basis—accordingly. The uncertainties are further exacerbated by error in the specification of the ambient into which the CMEs are cast, as discussed in conjunction with Figures 11, 15, and 16.

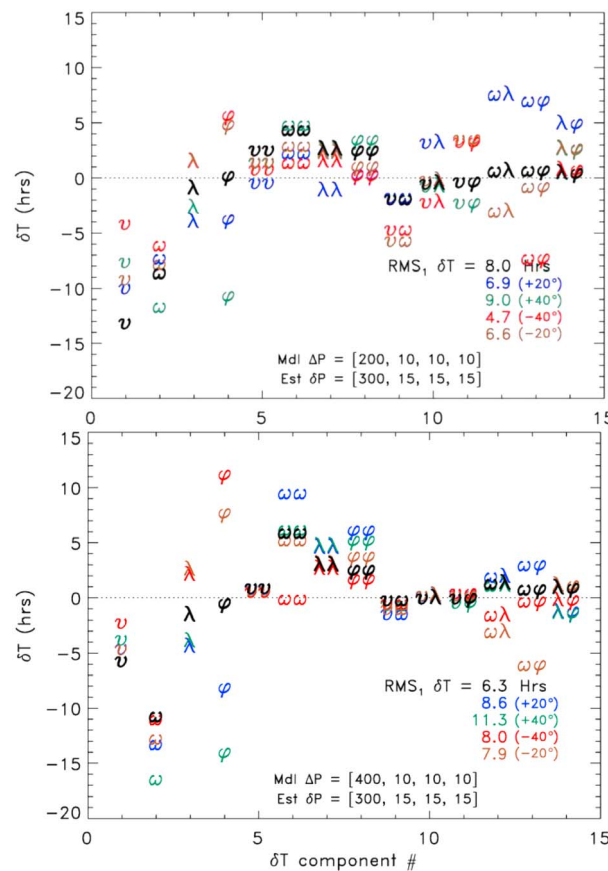As a final caveat to what we have presented, it must be borne in mind that the validity of the results applies strictly to single CMEs, and it remains to be seen whether these analyses can be extended to multiple, interacting CMEs. We would also expect the present results to hold—although with some quantitative adjustments—for CMEs with magnetic cloud as opposed to purely HD drivers, since the dynamics at the CME front should remain largely the same in the hypersonic solar wind.

7. Next Steps: Connection to the Real World

It has to be recognized that everything presented to this point is idealized, even if it is based quite rationally upon an operational model with some proven track record of success [e.g., *Millward et al.*, 2013]. Moreover, the strategies espoused above are based upon the assumption that the chaotic component of the real system must be small. Certainly, it is true that the most obvious signs of chaotic behavior are totally lacking in our modeling exercises, but compelling evidence of that proposal can only be had by direct comparison with the appropriate data. That is, Does convincing evidence for the fundamental trends and relationships described in our exposition actually exist in the observational record?

Figure 19. Taylor analysis diagrams for the (top) moderate-speed and (bottom) fast-speed 21-member ensembles of Figure 14, launched into the midst of the CIR front. The several colored symbol sets refer to Taylor analyses done at five different azimuths (0°, ±20°, and ±40°) along the equatorial plane, with black denoting that head-on to Earth (0° longitude). The fast-CME ensembles show more spread in the component distribution, and it is clear that CME size and pointing matter most for the faster CMEs. However, this figure is intended mainly to emphasize that in the presence of strong ambient structures the Taylor analysis implies that the overall forecast uncertainties become more complicated and increase in severity.

Fortunately, there is a body of data available that should go a long way toward answering this and other questions concerning the forecast strategies laid out in this paper. Namely, the SWPC record of official CME forecasts made in NRT since November 2011 can be mined to gauge the credibility of many of the assertions made in this work. Will, for example, a collection of real events exhibit the systematic behavior so prominent in the parameter jump versus transit time plots presented here? Even if that proves the case, can it be established that the plots and Taylor analysis developed here can in fact be used in real time to make forecasts more accurate, precise, and effective? Carefully drawn retrospective studies that may also include the full STEREO/LASCO CME record of CMEs not aimed toward Earth (i.e., not appearing in the SWPC record) should help pin down sources of error in this process and thus facilitate improvements in CME forecasting overall. It is only through comprehensive studies of real events that we can make sense of the sources of error and the overall credibility of this kind of forecast system.

Perusal of the Taylor component plots presented above raises a number of tactical questions. Given all the uncertainties and unknowns in this process, does it make any sense to include the cross derivatives in the analysis, and to what extent do even the second-order derivatives convey any actionable information? Does just the first-order set suffice, in the face of the observational unknowns? Limiting the analysis to only those terms would dramatically reduce the computational effort in producing NRT ensembles, but would anything useful be lost? And precisely what should guide how the model spreads Δv, etc., are specified in practice? For example, would $\Delta\lambda = \Delta\varphi = \omega_{1/2}$ make more sense than using some fixed value, as herein?

Finally, we speculate that adding two more parameters, α (related to the ambient speed) and μ (related to the CME mass) to the ensemble formalism may well prove more fruitful. Variations in the ambient speed would be assessed by adding two ensemble members having some spread $\Delta\alpha$ about the ambient value pertaining at the time the ensemble forecast is composed. Variations in the other parameter, μ (which can be adjusted via an Enlil internal parameter), would enable separating injected mass from geometric extent ω. The motivation is that the correlation between mass and angular width is only known on a statistical basis [e.g., *Vourlidas et al.*, 2010] and quite loosely at that. Also, we have at present no accurate way to gauge in NRT what the true mass may be for any given CME. While it appears CME mass and size may be in some sense related, we at least need to allow for separate contributions. Taylor analysis of ensembles covering the present first-order parameters require only nine members, and including both α and μ to first order would thus add only four more, for a total of 13 for an NRT run. Strategies such as this should be explored in the data-mining and NRT studies yet to come.

References

Arge, C. N., D. Odstrcil, V. J. Pizzo, and L. R. Mayer (2003), Improved method for specifying solar wind speed near the Sun, in *Proc. Solar Wind*, *AIP Conf. Ser.*, vol. 10, pp. 190–193, Pisa, Italy.

Arge, C. N., C. Henney, I. Gonzalez-Hernandez, W. Toussaint, J. Koller, and H. Godinez (2013), Modeling the corona and solar wind using ADAPT maps that include far-side, in *Solar Wind 13*, *AIP Conf. Ser.*, vol. 1539, pp. 11–14, Big Island, Hawaii.

Case, A. W., H. E. Spence, M. J. Owens, P. Riley, and D. Odstrcil (2008), Ambient solar wind's effect on ICME transit times, *Geophys. Res. Lett.*, *35*, L15105, doi:10.1029/2008GL034494.

DeYoung, D. S., and A. J. Hundhausen (1973), Simulation of driven flare-associated disturbances in the solar wind, *J. Geophys. Res.*, *78*, 3633–3642, doi:10.1029/JA078i019p03633.

Emmons, D., A. Acebal, A. Pulkkinen, A. Taktakishvili, P. MacNeice, and D. Odstrcil (2013), Ensemble forecasting of coronal mass ejections using the WSA-ENLIL with CONED Model, *Space Weather*, *11*, 95–106, doi:10.1002/swe.20019.

Godinez, H. C., and J. Koller (2012), Localized adaptive inflation in ensemble data assimilation for a radiation belt model, *Space Weather*, *10*, S08001, doi:10.1029/2012SW000767.

Gopalswamy, N., E. Aguilar-Rodriguez, S. Yashiro, S. Nunes, M. L. Kaiser, and R. A. Howard (2005), Type II radio bursts and energetic solar eruptions, *J. Geophys. Res.*, *110*, A12S07, doi:10.1029/2005JA011158.

Gubler, A., P. A. Catalán, and Y. Hayashi (2013), Probabilistic tsunami hazard assessment for near-field events, in *Coastal Dynamics 2013*, *7th Intl. Conf. Coastal Dyn.*, pp. 759–768, Bordeaux Univ., Bordeaux, France.

Hickmann, K. S., H. C. Godinez, C. J. Henney, and C. N. Arge (2015), Data assimilation in the ADAPT photospheric flux transport model, *Solar Phys.*, *290*, 1105–1118, doi:10.1007/s11207-015-0666-3.

Lee, C. O., C. N. Arge, D. Odstrcil, G. Millward, V. Pizzo, J. M. Quinn, and C. J. Henney (2013), Ensemble modeling of CME propagation, *Solar Phys.*, *285*, 349–368, doi:10.1007/s11207-012-9980-1.

Lee, C. O., C. N. Arge, D. Odstrcil, G. Millward, V. Pizzo, and N. Lugaz (2015), Ensemble modeling of successive halo CMEs: A case study, *Solar Phys.*, *290*, 1207–1229, doi:10.1007/s11207-015-0667-2.

Linker, J. A., Z. Mikic, P. Riley, C. Downs, R. Lionello, C. Henney, and C. N. Arge (2013), Coronal and heliospheric modeling using flux-evolved maps, in *Solar Wind 13*, *AIP Conf. Ser.*, vol. 1539, pp. 26–29, Big island, Hawaii.

Mays, M. L., et al. (2015), Ensemble modeling of CMEs using the WSA-ENLIL + Cone model, *Solar Phys.*, doi:10.1007/s11207-015-0692-1.

Millward, G., D. Biesecker, V. Pizzo, and C. A. de Koning (2013), An operational software tool for the analysis of coronagraph images: Determining CME parameters for input into the WSA-Enlil heliospheric model, *Space Weather*, *11*, 57–68, doi:10.1002/swe.200024.

Murray, S. A., E. M. Henley, D. R. Jackson, and S. L. Bruinsma (2015), Assessing the performance of thermospheric modeling with data assimilation throughout solar cycles 23 and 24, *Space Weather*, *13*, 220–232, doi:10.1002/2015SW001163.

Odstrcil, D. (2009), Numerical simulation of interplanetary disturbances, in *Proceedings of the Conf. 8-13 June, 2008 at the Westin Hotel, St. John, U.S. Virgin Islands, Numer. Model. Space Plasma Flows: ASTRONUM-2008, ASP Conf. Ser.*, vol. 406, edited by N. V. Pogorelov et al., p. 141, Astronomical Society of the Pacific, San Francisco, Calif.

Odstrcil, D., V. J. Pizzo, and C. N. Arge (2005), Propagation of the 12 May 1997 interplanetary CME in evolving solar wind structures, *J. Geophys. Res.*, *110*, A02106, doi:10.1029/2004JA010745.

Pizzo, V. (1980), A three-dimensional model of corotating streams in the solar wind: 2. Hydrodynamic streams, *J. Geophys. Res.*, *85*, 727–743, doi:10.1029/JA085iA02p00727.

Pizzo, V. (1982), A three-dimensional model of corotating streams in the solar wind: 3. Magnetohydrodynamic streams, *J. Geophys. Res.*, *87*, 4374–4394, doi:10.1029/JA087iA06p04374.

Pizzo, V. J. (1991), The evolution of corotating stream fronts near the ecliptic plane in the inner solar system: 2. Three-dimensional tilted-dipole fronts, *J. Geophys. Res.*, *96*, 5405–5420, doi:10.1029/91JA00155.

Pizzo, V., G. Millward, A. Parsons, D. Biesecker, S. Hill, and D. Odstrcil (2011), Wang-Sheeley-Arge-Enlil cone model transitions to operations, *Space Weather*, *9*, S03004, doi:10.1029/2011SW000663.

Riley, P., J. A. Linker, and Z. Mikic (2013), On the application of ensemble modeling techniques to improve ambient solar wind models, *J. Geophys. Res. Atmos.*, *118*, 600–607, doi:10.1002/jgra50156.

Shen, F., C. Shen, J. Zhang, P. Hess, Y. Wang, X. Feng, H. Cheng, and Y. Yang (2014), Evolution of the 12 July 2012 CME from the Sun to the Earth: Data-constrained three-dimensional MHD simulations, *J. Geophys. Res. Atmos.*, *119*, 7128–7141, doi:10.1002/2014JA020365.

Taktakishvili, A., P. MacNeice, and D. Odstrcil (2010), Model predictions of arrival of coronal mass ejections at Earth orbit, *Space Weather*, *8*, S06007, doi:10.1029/2009SW000543.

Vourlidas, A., R. A. Howard, E. Esfandiari, S. Patsourakos, S. Yashiro, and G. Michalek (2010), Comprehensive analysis of coronal mass ejection mass and energy properties over a full solar cycle, *Astrophys. J.*, *722*, 1522–1538, doi:10.1088/0004-637X/722/2/1522.